数理解析研究所講究録
第2030巻 2017年 47-54

A category of probability spaces and a conditional expectation functor*

Takanori Adachi (Ritsumeikan University)
and
Yoshihiro Ryu (Ritsumeikan University)

An advantage of using category theory is that it can visualize relations between different mathematical fields. Further, when we find a relation between different mathematical fields, it sometimes helps for developing a theory in a new direction. This fact motivates us to use category theory for studying probability theory.

One of the most prominent trials of applying category theory to probability theory so far is Lawvere and Giry's approach of formulating transition probabilities in a monad framework ([Lawvere, 1962], [Giry, 1982]). However, their approach is based on two categories, the category of measurable spaces (objects are measurable spaces and arrows are measurable maps) and the category of measurable spaces of a Polish space (objects are measurable spaces of a Polish space with a Borel σ-algebra and arrows are continuous maps) , not a category of probability spaces. Further, there are few trials of making categories consisting of all probability spaces due to a difficulty of finding an appropriate condition of their arrows.

Our approach is one of this simple-minded trials. We introduce a category **Prob** of all probability spaces in order to see a possible generalization of some classical tools in probability theory including conditional expectations. Actually, [Adachi, 2014] provides a simple category for formulating conditional expectations, but its objects and arrows are so limited that we cannot use it as a foundation of categorical probability theory.

Definition 1 (Category of Probability Spaces). A category **Prob** is the category whose objects are all probability spaces and the set of arrows between

*This work was supported by JSPS KAKENHI Grant Number 26330026.

them are defined by

$$\mathbf{Prob}(\bar{X},\bar{Y}) := \{f^- \mid f : \bar{Y} \to \bar{X} : \text{measurable with } \mathbb{P}_Y \circ f^{-1} \ll \mathbb{P}_X\},$$

where $\bar{X} := (X, \Sigma_X, \mathbb{P}_X), \bar{Y} := (Y, \Sigma_Y, \mathbb{P}_Y)$ and f^- is a symbol corresponding uniquely to a measurable function f.

We write $\bar{X} \xrightarrow{f^-} \bar{Y}$ in **Prob**, however, note that the arrow f^- has an opposite direction of the function f.

Now we are going to find a kind of conditional expectation in our category **Prob**. Let $f^- : \bar{X} \to \bar{Y}$ be an arbitrary arrow in **Prob**. For any $v \in \mathcal{L}^1(Y, \Sigma_Y, \mathbb{P}_Y)$, define a signed measure $v^* : \Sigma_Y \to \mathbf{R}$ as

$$v^*(B) := \int_B v \, d\mathbb{P}_Y \ (B \in \Sigma_Y).$$

Then, by the definition of arrow in **Prob**, a signed measure $v^* \circ f^{-1}$ on Σ_X is absolutely continuous relative to $\mathbb{P}_X$. So that, thanks to Radon-Nikodym theorem, we can find $E^{f^-}(v) \in \mathcal{L}^1(X, \Sigma_X, \mathbb{P}_X)$ as a Radon-Nikodym derivative of this signed measure $v^* \circ f^{-1}$ which is satisfying

$$\int_A E^{f^-}(v) \, d\mathbb{P}_X = \int_{f^{-1}(A)} v \, d\mathbb{P}_Y$$

for all $A \in \Sigma_X$. We call $E^{f^-}(v)$ a (version of) conditional expectation of v along f^-. This is a generalization of conditional expectation, because if $f = id_\Omega : (\Omega, \mathcal{F}, \mathbb{P}) \to (\Omega, \mathcal{G}, \mathbb{P})$ and $\mathcal{G} \subset \mathcal{F}$, then $E^{id_\Omega^-}(v)$ becomes a usual conditional expectation $\mathbb{E}(v|\mathcal{G})$. Further, we can think of an arrow f^- in **Prob** as a σ-algebra since the arrow $(\Omega, \mathcal{G}, \mathbb{P}) \xrightarrow{id_\Omega^-} (\Omega, \mathcal{F}, \mathbb{P})$ identifies a sub σ-algebra $\mathcal{G}$ of $\mathcal{F}$ as its domain.

Additionally, let $\sim_{\mathbb{P}}$ be $\mathbb{P}$-a.s. equivalence relation, then one can show

$$v_1 \sim_{\mathbb{P}_Y} v_2 \Rightarrow E^{f^-}(v_1) \sim_{\mathbb{P}_X} E^{f^-}(v_2),$$
$$E^{Id_X^-}(u) \sim_{\mathbb{P}_X} u,$$
$$E^{f^-}(E^{g^-}(w)) \sim_{\mathbb{P}_X} E^{g^- \circ f^-}(w)$$

for all $u \in \mathcal{L}^1(X, \Sigma_X, \mathbb{P}_X)$, $v_1, v_2 \in \mathcal{L}^1(Y, \Sigma_Y, \mathbb{P}_Y)$ and $w \in \mathcal{L}^1(Z, \Sigma_Z, \mathbb{P}_Z)$, where $\bar{X} \xrightarrow{f^-} \bar{Y} \xrightarrow{g^-} \bar{Z}$ and $\bar{X} \xrightarrow{Id_X^-} \bar{X}$. These imply well-definedness, identity preservility and composition preservility of the map $[v]_{\sim_{\mathbb{P}_Y}} \mapsto [E^{f^-}(v)]_{\sim_{\mathbb{P}_X}}$. So we have the first theorem:

Theorem 2 (Conditional Expectation Functor). *There exists a contravariant functor $\mathcal{E}$ from* **Prob** *to* **Set** *(the category of all sets and all functions) as following:*

$$\begin{tikzcd}
X & \bar{X} \arrow[r, mapsto, "\mathcal{E}"] \arrow[d, "f^-"'] & \mathcal{E}\bar{X} & := & L^1(X,\Sigma_X,\mathbb{P}_X) & \ni & {[E^{f^-}(v)]_{\sim_{\mathbb{P}_X}}} \\
Y \arrow[u, "f"] & \bar{Y} \arrow[r, mapsto, "\mathcal{E}"] & \mathcal{E}\bar{Y} \arrow[u, "\mathcal{E}f^-"'] & := & L^1(Y,\Sigma_Y,\mathbb{P}_Y) & \ni & {[v]_{\sim_{\mathbb{P}_Y}}.} \arrow[u, "\mathcal{E}f^-"']
\end{tikzcd}$$

We call $\mathcal{E}$ a ***conditional expectation functor****.*

Continually, we define a concept of measurability for our setting.

Definition 3 (Measurability). A random variable $v \in \mathcal{L}^\infty(Y,\Sigma_Y,\mathbb{P}_Y)$ is called f^--measurable if there exists $w \in \mathcal{L}^\infty(X,\Sigma_X,\mathbb{P}_X)$ such that $v \sim_{\mathbb{P}_Y} w \circ f$.

It seems natural because f^- is a "σ-algebra". More precisely, the arrow f^- identifies the σ-algebra $f^{-1}(\Sigma_X) = \sigma(f)$ and this definition is almost saying that v is $\sigma(f)$-measurable. Due to this definition, our second theorem is obtained.

Theorem 4 (Measurability). *Let u be an element of $\mathcal{L}^1(Y,\Sigma_Y,\mathbb{P}_Y)$ and v be a random variable in $\mathcal{L}^\infty(Y,\Sigma_Y,\mathbb{P}_Y)$, and assume that v is f^--measurable. Then we have*

$$E^{f^-}(v\cdot u) \sim_{\mathbb{P}_X} w \cdot E^{f^-}(u),$$

where $w \in \mathcal{L}^\infty(X,\Sigma_X,\mathbb{P}_X)$ is a random variable satisfying $v \sim_{\mathbb{P}_Y} w \circ f$.

A proof of theorem 4 can be obtained by using a usual step by step argument as the following: Firstly show it when w is an indicator function; Secondly show it if w is a simple function; Finally show it for general w.

This theorem shows that our "conditional expectation" still has a similar property about measurability.

Next definition is a modification of [Franz, 2003].

Definition 5 (Independence). We say $v \in \mathcal{L}^1(Y,\Sigma_Y,\mathbb{P}_Y)$ is independent of f^- if there exists a measure preserving map q which makes the following diagram commute:

$$\begin{tikzcd}
 & \bar{Y} \arrow[dl, "v"'] \arrow[d, "q"] \arrow[dr, "f"] & \\
(\mathbf{R},\mathcal{B},\mathbb{P}_Y\circ v^{-1}) & (\mathbf{R}\times X,\mathcal{B}\otimes\Sigma_X,(\mathbb{P}_Y\circ v^{-1})\otimes(\mathbb{P}_Y\circ f^{-1})) \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & \bar{X}.
\end{tikzcd}$$

By a straightforward calculation, we see that this definition means usual independence in the case of two σ-algebras. Indeed, by commutativity of the diagram, the map q must be equal to the map (v, f). Hence for all $C \in \mathcal{B}$ and $A \in \Sigma_X$,

$$\begin{aligned}
\mathbb{P}_Y(v^{-1}(C) \cap f^{-1}(A)) &= \mathbb{P}_Y(\{(v, f) \in C \times A\}) \\
&= \mathbb{P}_Y(q^{-1}(C \times A)) \\
&= (\mathbb{P}_Y \circ v^{-1}) \otimes (\mathbb{P}_Y \circ f^{-1})(C \times A) \\
&= \mathbb{P}_Y(v^{-1}(C)) \cdot \mathbb{P}_Y(f^{-1}(A)).
\end{aligned}$$

So σ-algebras $v^{-1}(\mathcal{B})$ and $f^{-1}(\Sigma_X)$ are independent under $\mathbb{P}_Y$. Furthermore, $v^{-1}(\mathcal{B})$ is nothing but $\sigma(v)$, and we think of $f^{-1}(\Sigma_X)$ as a given σ-algebra for conditional expectation. Thus the diagram just implies usual independence.

Finally, we encounter our last theorem.

Theorem 6 (Independence). *Let $v \in \mathcal{L}^1(Y, \Sigma_Y, \mathbb{P}_Y)$ be a random variable that is independent of f^-. Then we have,*

$$E^{f^-}(v) \sim_{\mathbb{P}_X} \mathbb{E}^{\mathbb{P}_Y}[v] E^{f^-}(1_Y).$$

When f is measure preserving, $E^{f^-}(1_Y) \sim_{\mathbb{P}_X} 1_X$, then the above formula turns to a well known formula of conditional expectation with independence, since $E^{f^-}(1_Y)$ is the Radon-Nikodym derivative $d(\mathbb{P}_Y \circ f^{-1})/d\mathbb{P}_X$.

Regarding proofs of theorem 6, one can prove this theorem by a usual method (using step functions and the dominated convergence theorem), but we want share a proof which is using commutative diagrams and functors. For this purpose, let us list some lemmas.

Lemma 7 (Functor **L**). *There exists a covariant functor* $\mathbf{L} : \mathbf{Prob} \to \mathbf{Set}$ *such that*

$$\begin{array}{ccccccc}
X & \bar{X} & \xmapsto{\mathbf{L}} & \mathbf{L}\bar{X} := L^\infty(X, \Sigma_X, \mathbb{P}_X) & \ni & [u]_{\sim_{\mathbb{P}_X}} \\
f \uparrow & f^- \downarrow & & \downarrow \mathbf{L}f^- & & \downarrow \mathbf{L}f^- \\
Y & \bar{Y} & \xmapsto{\mathbf{L}} & \mathbf{L}\bar{Y} := L^\infty(Y, \Sigma_Y, \mathbb{P}_Y) & \ni & [u \circ f]_{\sim_{\mathbb{P}_Y}}.
\end{array}$$

Sketch of Proof. Straightforward calculation with the definition of arrows in **Prob**, especially their absolute continuity. $\square$

Lemma 8 (Commutativity with Measure-Preserving). *If $f^- : \bar{X} \to \bar{Y}$ in* **Prob** *is measure-preserving, then we have $\mathcal{E}f^- \circ id_{\mathbf{L}\bar{Y}} \circ \mathbf{L}f^- = id_{\mathbf{L}\bar{X}}$, i.e. the diagram*

$$\begin{tikzcd} \mathbf{L}\bar{Y} \arrow[r, "id_{\mathbf{L}\bar{Y}}"] & \mathcal{E}\bar{Y} \arrow[d, "\mathcal{E}f^-"] \\ \mathbf{L}\bar{X} \arrow[u, "\mathbf{L}f^-"] \arrow[r, "id_{\mathbf{L}\bar{X}}"'] & \mathcal{E}\bar{X} \end{tikzcd}$$

commutes.

Proof. By theorem 4, for any $w \in \mathcal{L}^\infty(X, \Sigma_X, \mathbb{P}_X)$, we have

$$E^{f^-}(w \circ f) \sim_{\mathbb{P}_X} w \cdot E^{f^-}(1_Y).$$

However, since $E^{f^-}(1_Y)$ is nothing but a Radon-Nikodym derivative $d(\mathbb{P}_Y \circ f^{-1})/d\mathbb{P}_X$ and $f : (Y, \Sigma_Y, \mathbb{P}_Y) \to (X, \Sigma_X, \mathbb{P}_X)$ is measure-preserving, we see that

$$E^{f^-}(1_Y) \sim_{\mathbb{P}_X} \frac{d(\mathbb{P}_Y \circ f^{-1})}{d\mathbb{P}_X} \sim_{\mathbb{P}_X} \frac{d\mathbb{P}_X}{d\mathbb{P}_X} \sim_{\mathbb{P}_X} 1_X.$$

Thus $E^{f^-}(w \circ f) \sim_{\mathbb{P}_X} w$. In other words $\mathcal{E}f^- \circ id_{\mathbf{L}\bar{Y}} \circ \mathbf{L}f^- = id_{\mathbf{L}\bar{X}}$. □

Lemma 9 (Linearity). *Let $f^- : \bar{X} \to \bar{Y}$ be an arbitrary arrow in* **Prob**. *For all $u, v \in \mathcal{L}^1(Y, \Sigma_Y, \mathbb{P}_Y)$ and any $\alpha, \beta \in \mathbf{R}$,*

$$E^{f^-}(\alpha u + \beta v) \sim_{\mathbb{P}_X} \alpha E^{f^-}(u) + \beta E^{f^-}(v).$$

Sketch of Proof. Using a property of a Radon-Nikodym derivative with integral over subsets and linearity of integral. □

Lemma 10 (Monotone Convergence). *Let $f^- : \bar{X} \to \bar{Y}$ be an arbitrary arrow in* **Prob**. *Suppose that for any $n \in \mathbf{N}$, $v, v_n \in \mathcal{L}^1(Y, \Sigma_Y, \mathbb{P}_Y)$ and $0 \le v_n \uparrow v$ ($\mathbb{P}_Y$-a.s.). Then $0 \le E^{f^-}(v_n) \uparrow E^{f^-}(v)$, $\mathbb{P}_X$-almost surely.*

Sketch of Proof. Show that E^{f^-} is positive. Then put $u := \limsup_{n \to \infty} E^{f^-}(v_n)$ and prove this u is equal to $E^{f^-}(v)$ with the monotone convergence theorem. □

So we are ready to see a proof with diagrams.

Proof of Theorem 6. From the definition of independence, we have a commutative diagram

$$\begin{tikzcd} & \bar{Y} & \\ \bar{V} \arrow[ur, "v^-"] \arrow[r, "\pi_1^-"'] & \bar{V} \otimes \bar{X}^f \arrow[u, "q^-"] & \bar{X}, \arrow[ul, "f^-"'] \arrow[l, "\pi_2^-"] \end{tikzcd}$$

where $\bar{V} := (\mathbf{R}, \mathcal{B}, \mathbb{P}_Y \circ v^{-1})$ and $\bar{X}^f := (X, \Sigma_X, \mathbb{P}_Y \circ f^{-1})$. Then, because $\mathbf{L}$ and $\mathcal{E}$ are functors and lemma 8, each part of the diagram

$$\begin{tikzcd} & \mathbf{L}\bar{Y} \arrow[rr, "id_{\mathbf{L}\bar{Y}}"] & & \mathcal{E}\bar{Y} \arrow[d, "\mathcal{E}q^-"'] \arrow[dr, "\mathcal{E}f^-"] & \\ \mathbf{L}\bar{V} \arrow[ur, "\mathbf{L}v^-"] \arrow[r, "\mathbf{L}\pi_1^-"'] & \mathbf{L}(\bar{V} \otimes \bar{X}^f) \arrow[u, "\mathbf{L}q^-"'] \arrow[rr, "id_{\mathbf{L}(\bar{V}\otimes\bar{X}^f)}"'] & & \mathcal{E}(\bar{V} \otimes \bar{X}^f) & \mathcal{E}\bar{X} \arrow[l, "\mathcal{E}\pi_2^-"] \end{tikzcd}$$

commutes, hence the whole diagram also commutes. So that for any $[u]_{\sim_{\mathbb{P}_Y^v}} \in L^\infty(\mathbf{R}, \mathcal{B}, \mathbb{P}_Y \circ v^{-1})$, we obtain the following commutative diagram:

$$\begin{tikzcd} {[u \circ v]_{\sim_{\mathbb{P}_Y}}} \arrow[r, mapsto, "id_{\mathbf{L}\bar{Y}}"] & {[u \circ v]_{\sim_{\mathbb{P}_Y}}} \arrow[r, mapsto, "\mathcal{E}f^-"] & {[E^{f^-}(u \circ v)]_{\sim_{\mathbb{P}_X}}} \arrow[dd, equal] \\ {[u]_{\sim_{\mathbb{P}_Y^v}}} \arrow[u, mapsto, "\mathbf{L}v^-"] \arrow[d, mapsto, "\mathbf{L}\pi_1^-"'] & & \\ {[u \circ \pi_1]_{\sim_{\mathbb{P}_Y^v \otimes \mathbb{P}_Y^f}}} \arrow[r, mapsto, "id_{\mathbf{L}(\bar{V}\otimes\bar{X}^f)}"] & {[u \circ \pi_1]_{\sim_{\mathbb{P}_Y^v \otimes \mathbb{P}_Y^f}}} \arrow[r, mapsto, "\mathcal{E}\pi_2^-"] & {[E^{\pi_2^-}(u \circ \pi_1)]_{\sim_{\mathbb{P}_X}},} \end{tikzcd}$$

here $\mathbb{P}_Y^v := \mathbb{P}_Y \circ v^{-1}$ and $\mathbb{P}_Y^f := \mathbb{P}_Y \circ f^{-1}$. Thus $E^{f^-}(u \circ v) \sim_{\mathbb{P}_X} E^{\pi_2^-}(u \circ \pi_1)$.

However, for all $A \in \Sigma_X$,

$$\begin{aligned}
\int_A E^{\pi_2^-}(u \circ \pi_1)\, d\mathbb{P}_X &= \int_{\pi_2^{-1}(A)} u \circ \pi_1\, d(\mathbb{P}_Y^v \otimes \mathbb{P}_Y^f) \\
&= \int_{\mathbf{R}\times X} (u \circ \pi_1) \cdot (1_A \circ \pi_2)\, d(\mathbb{P}_Y^v \otimes \mathbb{P}_Y^f) \\
&= \mathbb{E}^{\mathbb{P}_Y^v \otimes \mathbb{P}_Y^f}[u \circ \pi_1] \cdot \mathbb{E}^{\mathbb{P}_Y^v \otimes \mathbb{P}_Y^f}[1_A \circ \pi_2] \\
&= \mathbb{E}^{\mathbb{P}_Y^v}[u] \cdot \mathbb{E}^{\mathbb{P}_Y^f}[1_A] \\
&= \mathbb{E}^{\mathbb{P}_Y}[u \circ v] \cdot \mathbb{E}^{\mathbb{P}_Y}[1_A \circ f] \\
&= \mathbb{E}^{\mathbb{P}_Y}[u \circ v] \int_{f^{-1}(A)} 1_Y\, d\mathbb{P}_Y \\
&= \int_A \mathbb{E}^{\mathbb{P}_Y}[u \circ v] \cdot E^{f^-}(1_Y)\, d\mathbb{P}_X.
\end{aligned}$$

Therefore

$$E^{f^-}(u \circ v) \sim_{\mathbb{P}_X} E^{\pi_2^-}(u \circ \pi_1) \sim_{\mathbb{P}_X} \mathbb{E}^{\mathbb{P}_Y}[u \circ v] E^{f^-}(1_Y).$$

Now put $u_n := id_{\mathbf{R}} \cdot 1_{[-n,n]}$, for any $n \in \mathbf{N}$. Then obviously $u_n \to id_{\mathbf{R}}$ as $n \to \infty$. So by lemma 9 and lemma 10, we obtain

$$\begin{aligned}
E^{f^-}(v) &\sim_{\mathbb{P}_X} \lim_{n\to\infty} E^{f^-}(u_n \circ v) \\
&\sim_{\mathbb{P}_X} \lim_{n\to\infty} \mathbb{E}^{\mathbb{P}_Y}[u_n \circ v] E^{f^-}(1_Y) \\
&\sim_{\mathbb{P}_X} \mathbb{E}^{\mathbb{P}_Y}[v] E^{f^-}(1_Y).
\end{aligned}$$

□

In conclusion, we provide a category **Prob** and a generalization of conditional expectation for this category which is called a conditional expectation functor $\mathcal{E}$. Also we show this generalized conditional expectation still has nice properties for measurability and independence. In addition, we give an unusual proof in probability theory which heavily uses the comutativity of diagrams and functors.

References

[Adachi, 2014] Adachi, T. (2014). Toward categorical risk measure theory. *Theory and Applications of Categories*, 29(14):389–405.

[Franz, 2003] Franz, U. (2003). What is stochastic independence? In *Quantum probability and White Noise Analysis, Non-commutativity, Infinite-dimensionality, and Probability at the Crossroads*, pages 254–274. World Sci. Publishing.

[Giry, 1982] Giry, M. (1982). A categorical approach to probability theory. In Banaschewski, B., editor, *Categorical Aspects of Topology and Analysis*, volume 915 of *Lecture Notes in Mathematics*, pages 68–85. Springer-Verlag.

[Lawvere, 1962] Lawvere, F. W. (1962). The category of probabilistic mappings. Preprint.